

SEQ ID NO: 1

SEQUENCE TYPE: nucleic acid

**TOPOLOGY:** linear

HYPOTHETICAL: No

10 ORIGINAL SOURCE:

TISSUE TYPE: brain

FEATURE KEY: 5' UTR

IDENTIFICATION METHOD: E

LOCATION: 76..2406

FEATURE KEY: 3' UTR

IDENTIFICATION METHOD: E

FEATURE KEY: polyA signal

LOCATION: 3978..4008

IDENTIFICATION METHOD: E

SEQUENCE DESCRIPTION:

5 GCCGAGGCCC GCGCAGTAGC GGTACTAAGT AGAGGCTGCT GGACGCGCCC CACCCGGCAC 60  
CAGGCGGAGC CAGAGATGCT TGCCAGGGCC GAGCGGCCCC GCCCGGGCCC CCGGCCGCCT 120  
CCGGTCTTTC CCTTCCCGCC GCCGCTGTCT CTGCTGCTGC TGCTGGCGAT ACTAAGCGCC 180  
CCGGTGTGCG GCCGCGTCCC CCGCTCAGTG CCCAGAACCT CGCTGCCCAT CTCCGAGGCT 240  
GACTCCTATC TCACCCGGTT TGCAGCGTCT CATACGTACA ATTACTCTGC TCTCCTTGTG 300  
10 GATCCTGCCT CCCACACACT TTACGTCGGT GCACGGGATA GCATCTTCGC TTAAACCCTC 360  
CCCTTCTCTG GGGAAAGACC CCGAAGGATC GACTGGATGG TACCTGAGAC TCACAGACAG 420  
AACTGCAGGA AGAAAGGCAA GAAAGAGGAC GAATGTCACA ATTTTATCCA GATTCTCGCC 480  
ATTGTCAATG CCTCTCACT CCTCAGTGC GGCACCTTCG CTTTGTATCC GAAGTGCGGG 540  
GTTATTGATG TGTCCAGTTT CCAGCAGGTT GAAAGACTTG AGAGCGGCCG GGGGAAATGT 600  
15 CCTTTGAGC CAGCTCAACG GTCAGCAGCT GTAATGGCTG GGGGCGTCCT CTACACCGCC 660  
ACTGTGAAGA ACTTCTGGG GACTGAGCCC ATCATCTCCC GAGCTGTGGG TCGAGCTGAG 720  
GACTGGATTC GAACAGAGAC CTTGTCATCC TGGCTTAATG CTCCAGCCTT TGTCGCAGCT 780  
ATGGTCTGA GCCCAGCTGA GTGGGGGGAT GAAGATGGAG ACGATGAAAT CTTTTTTTTC 840  
TTCACGGAGA CCTCCCGAGT GTTGGACTCC TATGAGCGCA TCAAGGTCCC AAGAGTGGCC 900  
20 CGAGTGTGTG CGGGGGACCT TGGGGGAGG AAGACCCTTC AGCAGAGATG GACGACGTTT 960  
CTGAAGGCTG ACCTGCTGTG CCCAGGGCCC GAGCATGCCG GGGCCTCCGG GGTCTGCAG 1020  
GCTATGGCAG AGCTTCGGCC TCAGCCTGGA GCGGGAACCC CCATCTTTTA TGGGATCTTT 1080  
TCCTCCAGT GGGAAAGAGC TGCCATCTCT GCTGTGTGTG CTTCCGACC CCAAGACATC 1140  
CGGGCAGTGC TGAATGGTCC CTTAGAGAG CTAAACATG ACTGCAACAG GGGACTGCCT 1200  
25 GTCATGGACA ACGAGGTGCC CCAGCCCAGA CCTGGAGAGT GATCGCCAA CAACATGAAG 1260

556090"08748260

	CTCCAGCAGT	TTGGATCCTC	ACTCTCCCTG	CCAGACCGCG	TGCTCACCTT	TATCAGAGAC	1320
	CACCTCTCA	TGGACAGGCC	CGTGTTCCTG	GCTGACGGCC	GCCCCCTGCT	GGTCACTACA	1380
	GATACAGCCT	ATCTCAGAGT	CGTGGCCAC	AGGGTGACCA	GCCTCTCAGG	GAAAGAATAT	1440
	GACGTGCTCT	ACCTGGGGAC	AGAGGATGGA	CACCTCCACC	GGGCTGTGCG	CATTGAGCT	1500
5	CAGCTCAGTG	TCTTGGAGGA	TCTGGCCTTG	TTCCCAGAAC	CACAGCCGGT	TGAGAGCATG	1560
	AAATTGTACC	ACGATTGGCT	CCTGGTGGGC	TCCCATACTG	AGGTGACACA	AGTGAACACC	1620
	AGCAACTGTG	GCCGTCTCCA	GAGCTGCTCG	GAGTGTATCC	TGGCCCAAGG	CCCCGTGTGC	1680
	GCCTGGAGCT	TCCGGCTTGA	TGCTTGTGTG	GCCACGCCG	GCGAGCACCG	CGGGATGGTT	1740
	CAAGATATAG	AGTCAGCGGA	TGTCTCTTCT	TTGTGTCAA	AAGAACCTGG	AGAACATCCC	1800
10	GTAGTGTTTG	AAGTTCCGGT	GGCTACTGTG	GGCCACGTGG	TCCTGCCATG	TTCCCCAGT	1860
	TCTGCCTGGG	CATCCTGTGT	GTGGCACCAG	CCCAGTGGAG	TGACTGCGCT	CACTCCCCGG	1920
	AGGGATGGAC	TAGAGGTGGT	GGTGACCCCA	GGGGCCTATG	GGGCTTATGC	TTGCGAGTGT	1980
	CAGGAGGGTG	GAGCCGCCCG	CGTGGTGGCT	GCTTATAGCT	TGGTGTGGGG	CAGCCAGCGG	2040
	GGACCCTCAA	ACCGGGCCCA	CACCGTTGTG	GGGGCTGGAT	TGGTTGGCTT	TCTCCTGGGT	2100
15	GTTCTTGCAG	CATCCCTCAC	TCTCCTCCTG	ATTGGTCGCC	GTCAGCAGCG	TCGGCGACAG	2160
	AGGGAGCTTC	TAGCTAGAGA	CAAGGTGGGC	TTAGATCTGG	GGGCTCCACC	TTCTGGGACC	2220
	ACAAGCTATA	GTCAGGACCC	TCCCTCTCCT	TCGCCTGAAG	ATGAACGGCT	GCCCCTGGCC	2280
	CTGGGTAAGC	GGGGCAGTGG	TTTTGGTGGC	TTCCCTCCAC	CCTTCTGCT	GGATTCTTGC	2340
	CCAAGCCCAG	CCCACATCCG	GCTCACTGGG	GCGCCTCTAG	CCACGTGTGA	TGAGACCTCC	2400
20	ATCTAAAGCC	GGGAAAATG	ACTGCCAGCC	ATGAGCAGTC	TCTGGAATA	GTGGCTACCA	2460
	AGACCATGAT	CATGGCTGCT	CCTTTCTCTT	GGAGTCTGTG	TGTTACACA	TTAGTGTCTG	2520
	TCCTCTGGAC	CTGGACCTGG	CCTTTGCCCA	GATTCTGAT	TCTCATGAGA	GATCAACCCT	2580
	GTAACCTTCT	GCGATGGCT	CTTGTCTTGG	GCCCATCAGC	TTGTGGGGTG	GAGTAAGGAC	2640
	ATAGGCCCCG	GAAAGGCAAT	CAGTGTGGAG	GAGTTGGGG	CGTGTGTGCC	CTGCGTCTTT	2700
25	GTGGTGGCTG	TATGATTTCC	CAGTCTGCTG	ACTCTGGGGA	GCGCATGATC	CCCTGACTGC	2760

SEQ ID NO: 2

25 SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: No

5 ANTI-SENSE: No

ORIGINAL SOURCE:

ORGANISM: rat (*Rattus norvegicus*)

TISSUE TYPE: brain

FEATURE:

10 FEATURE KEY: CDS

LOCATION: 1..2331

IDENTIFICATION METHOD: E

SEQUENCE DESCRIPTION:

ATGCTTGCCA GGGCCGAGCG GCCCCGCCCG GGC0000GGC CGCCTCCGGT CTTTCCCTTC 60  
15 CCGCCGCCGC TGTCGCTGCT GCTGCTGCTG GCGATACTAA GCGCCCCGGT GTGCGGCCGC 120  
GTCCCCCGCT CAGTGCCCAG AACCTCGCTG CCCATCTCCG AGGCTGACTC CTATCTCACC 180  
CGGTTTGAGC CGTCTCATAC GTACAATTAC TCTGCTCTCC TTGTGGATCC TGCCTCCCAC 240  
ACACTTTACG TCGGTGCACG GGATAGCATC TTCGCTTTAA CCCTCCCCTT CTCTGGGGAA 300  
AGACCCCGAA GGATCGACTG GATGGTACCT GAGACTCACA GACAGAACTG CAGGAAGAAA 360  
20 GGCAAGAAAG AGGACGAATG TCACAATTTT ATCCAGATTC TCGCCATTGT CAATGCCTCT 420  
CACCTCTCA CGTGCGGCAC CTTGCTTTT GATCCGAAGT GCGGGGTAT TGATGTGTCC 480  
AGTTTCCAGC AGGTTGAAAG A0TTGAGAGC GGCCGGGGGA AATGTCCTTT TGAGCCAGCT 540  
CAACGGTCAG CAGCTGTAAT G0CTGGGGGC GTCCTCTACA CCGCCACTGT GAAGAACTTC 600  
CTGGGGACTG AGCCCATCAT CTCCCGAGCT GTGGGTCGAG CTGAGGACTG GATTCCAACA 660  
25 GAGACCTTGT CATCCTGGCT TAATGCTCCA GCCTTTGT0G CAGCTATGGT CCTGAGCCCA 720

a

09284180 060999

5 6 6 0 9 0 . 0 8 1 8 2 6 0

	GCTGAGTGGG	GGGATGAAGA	TGGAGACGAT	GAAATCTTTT	TTTCTTCAC	GGAGACCTCC	780
	CGAGTGTGG	ACTCCTATGA	GCGCATCAAG	GTCCCAAGAG	TGGCCGAGT	GTGTGCGGGG	840
	GACCTTGGGG	GCAGGAAGAC	CCTTCAGCAG	AGATGGACGA	CGTTTCTGAA	GGCTGACCTG	900
	CTGTGCCCAG	GGCCCCGAGCA	TGGCCGGGCC	TCCGGGGTTC	TGCAGGCTAT	GGCAGAGCTT	960
5	CGGCCTCAGC	CTGGAGCGGG	AACCCCATC	TTTTATGGGA	TCTTTTCCTC	CCAGTGGGAA	1020
	GGAGCTGCCA	TCTCTGCTGT	GTGTGCCTTC	CGACCCCAAG	ACATCCGGGC	AGTGCTGAAT	1080
	GGTCCCTTTA	GAGAGCTAAA	ACATGACTGC	AACAGGGGAC	TGCCTGTCA	GGACAACGAG	1140
	GTGCCCCAGC	CCAGACCTGG	AGAGTGCATC	GCCAACAACA	TGAAGCTCCA	GCAGTTTGGA	1200
	TCCTCACTCT	CCCTGCCAGA	CCGCGTGCTC	ACCTTTATCA	GAGACCACCC	TTCATGGAC	1260
10	AGGCCCGTGT	TCCCGGCTGA	CGGCCGCCCC	CTGCTGGTCA	CTAGAGATAC	AGCCTATCTC	1320
	AGAGTCGTGG	CCCACAGGGT	GACCAGCCTC	TCAGGGAAAG	AATATGACGT	GCTCTACCTG	1380
	GGGACAGAGG	ATGGACACCT	CCACGGGGCT	GTGCGCATTG	GAGCTCAGCT	CAGTGTCTTG	1440
	GAGGATCTGG	CCTTGTTCCC	AGAACCACAG	CCGGTTGAGA	GCATGAAATT	GTACCACGAT	1500
	TGGCTCCTGG	TGGGCTCCCA	TACTGAGGTG	ACACAAGTGA	ACACCAGCAA	CTGTGGCCGT	1560
15	CTCCAGAGCT	GCTCGGAGTG	TATCCTGGCC	CAGGACCCCG	TGTGCGCCTG	GAGCTTCCGG	1620
	CTTGATGCTT	GTGTGGCCCA	CGCCGGCGAG	CACCGCGGGA	TGGTTCAAGA	TATAGAGTCA	1680
	GCGGATGTCT	CTTCTTTGTG	TCCAAAAGAA	CCTGGAGAAC	ATCCCGTAGT	GTTTGAAGTT	1740
	CCGGTGGCTA	CTGTGGGCCA	CGTGGTCCTG	CCATGTTCCC	CCAGTTCTGC	CTGGGCATCC	1800
	TGTGTGTGGC	ACCAGCCCAG	TGGAGTGA	CGGCTCACTC	CCCGGAGGGA	TGGA	1860
20	GTGGTGGTGA	CCCCAGGGGC	CATGGGGGCT	TATGCTTGCG	AGTGCAGGA	GGGTGGAGCC	1920
	GCCCGCGTGG	TGGCTGCTTA	TAGCTTGGTG	TGGGGCAGCC	AGCGGGGACC	CTCAAACCGG	1980
	GCCCACACCG	TTGTGGGGGC	TGGATTGGTT	GGCTTTCTCC	TGGGTGTTCT	TGCAGCATCC	2040
	CTCACTCTCC	TCCTGATTGG	TCGCCGTCAG	CAGCGTCGGC	GACAGAGGGA	GCTTCTAGCT	2100
	AGAGACAAGG	TGGGCTTAGA	TCTGGGGGCT	CCACCTTCTG	GGACCACAAG	CTATAGTCAG	2160
25	GACCCTCCCT	CTCCTTCGCC	TGAAGATCAA	CGGCTGCCCC	TGGCCCTGGG	TAAGCGGGGC	2220

ca

AGTGGTTTTG GTGGCTTCCC TCCACCTTC CTGCTGGATT CTTGCCCAAG CCCAGCCCAC 2280  
ATCCGGCTCA CTGGGGCGCC TCTAGCCACG TGTGATGAGA CCTCCATCTA A 2331

SEQ ID NO: 3

5 SEQUENCE LENGTH: 776 amino acids

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

10 ORGANISM: rat (*Rattus norvegicus*)

TISSUE TYPE: brain

FEATURE:

FEATURE KEY: peptide

LOCATION: 1..776

15 IDENTIFICATION METHOD: P

SEQUENCE DESCRIPTION:

Met Leu Ala Arg Ala Glu Arg Pro Arg Pro Gly Pro Arg Pro Pro Pro

1 5 10 15

Val Phe Pro Phe Pro Pro Pro Leu Ser Leu Leu Leu Leu Ala Ile

20 20 25 30

Leu Ser Ala Pro Val Cys Gly Arg Val Pro Arg Ser Val Pro Arg Thr

35 40 45

Ser Leu Pro Ile Ser Glu Ala Asp Ser Tyr Leu Thr Arg Phe Ala Ala

50 55 60

25 Ser His Thr Tyr Asn Tyr Ser Ala Leu Leu Val Asp Pro Ala Ser His

09284180 060999

1a

	65		70		75		80									
	Thr	Leu	Tyr	Val	Gly	Ala	Arg	Asp	Ser	Ile	Phe	Ala	Leu	Thr	Leu	Pro
				85						90					95	
	Phe	Ser	Gly	Glu	Arg	Pro	Arg	Arg	Ile	Asp	Trp	Met	Val	Pro	Glu	Thr
5				100					105						110	
	His	Arg	Gln	Asn	Cys	Arg	Lys	Lys	Gly	Lys	Lys	Glu	Asp	Glu	Cys	His
				115					120					125		
	Asn	Phe	Ile	Gln	Ile	Leu	Ala	Ile	Val	Asn	Ala	Ser	His	Leu	Leu	Thr
				130					135					140		
10	Cys	Gly	Thr	Phe	Ala	Phe	Asp	Pro	Lys	Cys	Gly	Val	Ile	Asp	Val	Ser
				145					150					155		160
	Ser	Phe	Gln	Gln	Val	Glu	Arg	Leu	Glu	Ser	Gly	Arg	Gly	Lys	Cys	Pro
				165					170					175		
	Phe	Glu	Pro	Ala	Gln	Arg	Ser	Ala	Ala	Val	Met	Ala	Gly	Gly	Val	Leu
15				180					185					190		
	Tyr	Thr	Ala	Thr	Val	Lys	Asn	Phe	Leu	Gly	Thr	Glu	Pro	Ile	Ile	Ser
				195					200					205		
	Arg	Ala	Val	Gly	Arg	Ala	Glu	Asp	Trp	Ile	Arg	Thr	Glu	Thr	Leu	Ser
				210					215					220		
20	Ser	Trp	Leu	Asn	Ala	Pro	Ala	Phe	Val	Ala	Ala	Met	Val	Leu	Ser	Pro
				225					230					235		240
	Ala	Glu	Trp	Gly	Asp	Glu	Asp	Gly	Asp	Asp	Glu	Ile	Phe	Phe	Phe	Phe
				245					250					255		
	Thr	Glu	Thr	Ser	Arg	Val	Leu	Asp	Ser	Tyr	Glu	Arg	Ile	Lys	Val	Pro
25				260					265					270		



Arg Val Ala Arg Val Cys Ala Gly Asp Leu Gly Gly Arg Lys Thr Leu

275

280

285

Gln Gln Arg Trp Thr Thr Phe Leu Lys Ala Asp Leu Leu Cys Pro Gly

290

295

300

5 Pro Glu His Gly Arg Ala Ser Gly Val Leu Gln Ala Met Ala Glu Leu

305

310

315

320

Arg Pro Gln Pro Gly Ala Gly Thr Pro Ile Phe Tyr Gly Ile Phe Ser

325

330

335

Ser Gln Trp Glu Gly Ala Ala Ile Ser Ala Val Cys Ala Phe Arg Pro

10

340

345

350

Gln Asp Ile Arg Ala Val Leu Asn Gly Pro Phe Arg Glu Leu Lys His

355

360

365

Asp Cys Asn Arg Gly Leu Pro Val Met Asp Asn Glu Val Pro Gln Pro

370

375

380

15 Arg Pro Gly Glu Cys Ile Ala Asn Asn Met Lys Leu Gln Gln Phe Gly

385

390

395

400

Ser Ser Leu Ser Leu Pro Asp Arg Val Leu Thr Phe Ile Arg Asp His

405

410

415

Pro Leu Met Asp Arg Pro Val Phe Pro Ala Asp Gly Arg Pro Leu Leu

20

420

425

430

Val Thr Thr Asp Thr Ala Tyr Leu Arg Val Val Ala His Arg Val Thr

435

440

445

Ser Leu Ser Gly Lys Glu Tyr Asp Val Leu Tyr Leu Gly Thr Glu Asp

450

455

460

25 Gly His Leu His Arg Ala Val Arg Ile Gly Ala Gln Leu Ser Val Leu

09234180-060999

*Handwritten signature*

09284180-050999

	465	470	475	480
	Glu Asp Leu Ala Leu Phe Pro Glu Pro Gln Pro Val Glu Ser Met Lys			
	485	490	495	
	Leu Tyr His Asp Trp Leu Leu Val Gly Ser His Thr Glu Val Thr Gln			
5	500	505	510	
	Val Asn Thr Ser Asn Cys Gly Arg Leu Gln Ser Cys Ser Glu Cys Ile			
	515	520	525	
	Leu Ala Gln Asp Pro Val Cys Ala Trp Ser Phe Arg Leu Asp Ala Cys			
	530	535	540	
10	Val Ala His Ala Gly Glu His Arg Gly Met Val Gln Asp Ile Glu Ser			
	545	550	555	560
	Ala Asp Val Ser Ser Leu Cys Pro Lys Glu Pro Gly Glu His Pro Val			
	565	570	575	
	Val Phe Glu Val Pro Val Ala Thr Val Gly His Val Val Leu Pro Cys			
15	580	585	590	
	Ser Pro Ser Ser Ala Trp Ala Ser Cys Val Trp His Gln Pro Ser Gly			
	595	600	605	
	Val Thr Ala Leu Thr Pro Arg Arg Asp Gly Leu Glu Val Val Val Thr			
	610	615	620	
20	Pro Gly Ala Met Gly Ala Tyr Ala Cys Glu Cys Gln Glu Gly Gly Ala			
	625	630	635	640
	Ala Arg Val Val Ala Ala Tyr Ser Leu Val Trp Gly Ser Gln Arg Gly			
	645	650	655	
	Pro Ser Asn Arg Ala His Thr Val Val Gly Ala Gly Leu Val Gly Phe			
25	660	665	670	

*Handwritten signature*

Leu Leu Gly Val Leu Ala Ala Ser Leu Thr Leu Leu Leu Ile Gly Arg

675

680

685

Arg Gln Gln Arg Arg Arg Gln Arg Glu Leu Leu Ala Arg Asp Lys Val

690

695

700

5 Gly Leu Asp Leu Gly Ala Pro Pro Ser Gly Thr Thr Ser Tyr Ser Gln

705

710

715

720

Asp Pro Pro Ser Pro Ser Pro Glu Asp Glu Arg Leu Pro Leu Ala Leu

725

730

735

Gly Lys Arg Gly Ser Gly Phe Gly Gly Phe Pro Pro Pro Phe Leu Leu

10

740

745

750

Asp Ser Cys Pro Ser Pro Ala His Ile Arg Leu Thr Gly Ala Pro Leu

755

760

765

Ala Thr Cys Asp Glu Thr Ser Ile

770

775

15

SEQ ID NO: 4

SEQUENCE LENGTH: 2315 base pairs

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

20 TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: No

ANTI-SENSE: No

ORIGINAL SOURCE:

25 ORGANISM: human (Homo sapiens)

a

09284180.060999

TISSUE TYPE: brain

FEATURE:

FEATURE KEY: CDS

LOCATION: 1..1764

5 IDENTIFICATION METHOD: E

FEATURE KEY: 3' UTR

LOCATION: 1765..2315

IDENTIFICATION METHOD: E

10 SEQUENCE DESCRIPTION:

GGGGGTGTCC TCTATGCTGC CACTGTGAAA AACTACCTGG GGACGGAGCC AATTATCACC 60

AGAGCAGTGG GTCGTGCCGA GGACTGGATT CGGACAGATA CCTTGCCTTC CTGGCTGAAC 120

GCCCCAGCCT TTGTCGCAGC CGTGGCCTTG AGCCCAGCCG AATGGGGGGA TGAAGATGGA 180

GACGACGAAA TCTACTTCTT CTTTACGGAG ACTTCCCGAG CATTTGACTC ATACGAGCGC 240

15 ATTAAAGTCC CACGGGTGGC CCGTGTGTGT GCGGGGGACC TCGGGGGCCG GAAGACCCTC 300

CAGCAGAGAT GGACGACGTT TTTGAAAGCT GACCTGCTCT GTCCAGGGCC TGAGCATGGC 360

CGGGCCTCCA GTGTCCTGCA GGATGTTGCT GTGCTTCGAC CTGAGCTTGG GGCAGGGACT 420

CCCATCTTTT ATGGCATCTT TTCTTCCCAG TGGGAGGGGG CTAATATCTC TGCTGTCTGT 480

GCCTTCCGAC CACAAGACAT TCGGACAGTG CTGAATGGTC CCTTCAGAGA ACTAAAACAT 540

20 GACTGCAACA GAGGACTGCC TGTCGTGGAC AATGATGTGC CCCAGCCCAG ACCTGGAGAG 600

TGCATCACCA ACAACATGAA GCTCCGGGAC TTTGGCTCAT CTCTCTCCCT GCCTGACCGC 660

GTACTCACCT TCATCCGGGA CCACCCACTC ATGGACAGGC CAGTGTTTCC AGCTGATGGC 720

CACCCCTGC TGGTCACTAC AGATACAGCC TATCTCAGAG TCGTGGCCCA CAGGGTGACC 780

AGCCTCTCAG GGAAAGAGTA TGATGTGCTC TACCTGGGGA CAGAGGATGG ACACCTCCAC 840

25 CGAGCAGTGC GGATCGGAGC TCAGCTCAGC GTTCTTGAAG ATCTGGCCTT ATTCCCAGAG 900

0084180 050990

Q

00284180 060909

CCACAGCCAG TTGAGAACAT GAAATTGTAC CACAGCTGGC TCCTGGTTGG CTCCCGTACT 960  
GAGGTGACAC AAGTGAATAC AACCAACTGT GGCCGTCTCC AGAGCTGCTC AGAGTCCATC 1020  
CTGGCCCAGG ACCCAGTCTG TGCCTGGAGC TTCCGGCTGG ATGAGTGTGT GGCCCATGCC 1080  
GGGGAGCACC GAGGGTTGGT CCAAGACATA GAGTCAGCAG ATGTCTCCTC TTTGTGTCCT 1140  
5 AAAGAGCCTG GAGAACGTCC AGTAGTGTTT GAAGTTCCCG TGGCTACAGC TGCGCATGTG 1200  
GTCTTGCCAT GTTCTCCAAG CTCAGCATGG GCATCCTGTG TGTGGCAGCA GCCCAGTGGA 1260  
GTGACTGCAC TCACCCCCCG GCGGGATGGA CTGGAGGTGG TGGTGACCCC AGGGGCCATG 1320  
GGCGCTTATG CCTGTGAATG TCAGGAGGGT GGGGCAGCCC ATGTGGTAGC AGCTTACAGC 1380  
TTGGTATGGG GCAGCCAGCG AGATGCTCCG AGCCGGGCCC ACACAGTGGG GGCGGGACTG 1440  
10 GCTGGCTTCT TCTTGGGGAT TCTCGCAGCA TCCCTGACTC TCATTCTGAT TGGTCGGCGT 1500  
CAGCAGCGAC GGCAGACAGAG GGAACCTCTG GCTAGAGACA AGGTGGGCCT GGACCTGGGG 1560  
GCTCCACCTT CTGGGACCAC AAGCTACAGC CAAGACCCTC CCTCCCCCTC TCCTGAAGAT 1620  
GAGCGGTTGC CGCTGGCCCT GGCCAAGAGG GGCAGTGGCT TTGGTGGATT CTCACCACCC 1680  
TTCCTGCTTG ATCCTTGCCC AAGCCCAGCC CACATTGCGC TAACTGGGGC TCCTCTAGCC 1740  
15 ACATGTGATG AAACATCCAT CTAGAGCTGG GCAAATGACC ACTAGTGTAT AAGTGATCAC 1800  
TGGAACGGAG TGACCACTGA GATGCTGGGG GTCAGTGGGC CTGGAAGACC ATCCCAGCCT 1860  
CTGAGTTCTC TTTGAGTATG AGTGATTACT TGGATTTTAG TATCTGTTCT CTCTGAGCCT 1920  
GGATGGGCTT GGGGCCAGAC CTTTGCTGA TTCCTGATTC CCATGAGAAA TCAGAACTGC 1980  
TTTCTGCAGC AAATCAGGGC TTCCCCCTAA CATCTGAACT CCTGTAAACC TTCATCCCTG 2040  
20 GCCCCCTATC TTGGGCCCAT TAGTTTTGGG GATGGGGCAC AGGGCATAGC TATGACTTTG 2100  
CTTCTGGTT GGAGCCTGGC CGGAAGGAAG AGCCCTGGAG GTGGTTGGGG GCAAATGTGC 2160  
CCTGAGTCCT TGGGGTGGTT CTGCTTATTC TTCAAGTTTA TCTGAATCTG TGGGGAGTGC 2220  
ATGATCCCCA TGTTGCAATA TGGAGTCTCT GCCCTGAGAT CTTCCCCATC TCAGTTTTCC 2280  
TTCCATGAAA GAGTACGTGT AAATACATAG TGTTTC 2315

SEQ ID NO: 5

SEQUENCE LENGTH: 1761 base pairs

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

5 TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: No

ANTI-SENSE: No

ORIGINAL SOURCE:

10 ORGANISM: human (Homo sapiens)

TISSUE TYPE: brain

FEATURE:

FEATURE KEY: CDS

LOCATION: 1..1761

15 IDENTIFICATION METHOD: E

SEQUENCE DESCRIPTION:

GGGGGTGTCC TCTATGCTGC CACTGTGAAA AACTACCTGG GGACGGAGCC AATTATCACC 60  
AGAGCAGTGG GTCGTGCCGA GGACTGGATT CGGACAGATA CCTTGCCTTC CTGGCTGAAC 120  
GCCCCAGCCT TTGTCGCAGC CGTGGCCTTG AGCCCAGCCG AATGGGGGGA TGAAGATGGA 180  
20 GACGACGAAA TCTACTTCTT CTTTACGGAG ACTTCCCGAG CATTTGACTC ATACGAGCGC 240  
ATTAAAGTCC CACGGGTGGC CCGTGTGTGT GCGGGGGACC TCGGGGGCCG GAAGACCCTC 300  
CAGCAGAGAT GGACGACGTT TTTGAAAGCT GACCTGCTCT GTCCAGGGCC TGAGCATGGC 360  
CGGGCCTCCA GTGTCCTGCA GGATGTTGCT GTGCTTCGAC CTGAGCTTGG GGCAGGGACT 420  
CCCATCTTTT ATGGCATCTT TTCTTCCAG TGGGAGGGGG CTAATATCTC TGCTGTCTGT 480  
25 GCCTTCCGAC CACAAGACAT TCGGACAGTG CTGAATGGTC CCTTCAGAGA ACTAAAACAT 540

09284180.060999

GACTGCAACA GAGGACTGCC TGTCGTGGAC AATGATGTGC CCCAGCCCAG ACCTGGAGAG 600  
 TGCATCACCA ACAACATGAA GCTCCGGCAC TTTGGCTCAT CTCTCTCCCT GCCTGACCGC 660  
 GTACTCACCT TCATCCGGGA CCACCCACTC ATGGACAGGC CAGTGTTCCT AGCTGATGGC 720  
 CACCCCTGCT TGGTCACTAC AGATACAGCC TATCTCAGAG TCGTGGCCCA CAGGGTGACC 780  
 5 AGCCTCTCAG GGAAAGAGTA TGATGTGCTC TACCTGGGGA CAGAGGATGG ACACCTCCAC 840  
 CGAGCAGTGC GGATCGGAGC TCAGCTCAGC GTTCTTGAAG ATCTGGCCTT ATTCCCAGAG 900  
 CCACAGCCAG TTGAGAACAT GAAATTGTAC CACAGCTGGC TCCTGGTTGG CTCCCGTACT 960  
 GAGGTGACAC AAGTGAATAC AACCAACTGT GGCCGTCTCC AGAGCTGCTC AGAGTGCATC 1020  
 CTGGCCCAGG ACCCAGTCTG TGCCTGGAGC TTCCGGCTGG ATGAGTGTGT GGCCCATGCC 1080  
 10 GGGGAGCACC GAGGGTTGGT CCAAGACATA GAGTCAGCAG ATGTCTCCTC TTTGTGTCCT 1140  
 AAAGAGCCTG GAGAACGTCC AGTAGTGTTT GAAGTTCCCG TGGCTACAGC TGCGCATGTG 1200  
 GTCTTGCCAT GTTCTCCAAG CTCAGCATGG GCATCCTGTG TGTGGCACCA GCCCAGTGGA 1260  
 GTGACTGCAC TCACCCCCCG GCGGGATGGA CTGGAGGTGG TGGTGACCCC AGGGGCCATG 1320  
 GGCGCTTATG CCTGTGAATG TCAGGAGGGT GGGGCAGCCC ATGTGGTAGC AGCTTACAGC 1380  
 15 TTGGTATGGG GCAGCCAGCG AGATGCTCCG AGCCGGGCCC ACACAGTGGG GGCGGGACTG 1440  
 GCTGGCTTCT TCTTGGGGAT TCTCGCAGCA TCCCTGACTC TCATTCTGAT TGGTCGGCGT 1500  
 CAGCAGCGAC GGCGACAGAG GGAAGTTCTG GCTAGAGACA AGGTGGGCCT GGACCTGGGG 1560  
 GCTCCACCTT CTGGGACCAC AAGCTACAGC CAAGACCCTC CCTCCCCCTC TCCTGAAGAT 1620  
 GAGCGGTTGC CGCTGGCCCT GGCCAAGAGG GGCAGTGGCT TTGGTGGATT CTCACCACCC 1680  
 20 TTCCTGCTTG ATCCTTGCCC AAGCCCAGCC CACATTGGGC TAACTGGGGC TCCTCTAGCC 1740  
 ACATGTGATG AAACATCCAT C 1761

SEQ ID NO: 6

SEQUENCE LENGTH: 587 amino acids

25 SEQUENCE TYPE: amino acid



092841801050999

2





Glu Val Thr Gln Val Asn Thr Thr Asn Cys Gly Arg Leu Gln Ser Cys

325

330

335

Ser Glu Cys Ile Leu Ala Gln Asp Pro Val Cys Ala Trp Ser Phe Arg

340

345

350

5 Leu Asp Glu Cys Val Ala His Ala Gly Glu His Arg Gly Leu Val Gln

355

360

365

Asp Ile Glu Ser Ala Asp Val Ser Ser Leu Cys Pro Lys Glu Pro Gly

370

375

380

Glu Arg Pro Val Val Phe Glu Val Pro Val Ala Thr Ala Ala His Val

10 385

390

395

400

Val Leu Pro Cys Ser Pro Ser Ser Ala Trp Ala Ser Cys Val Trp His

405

410

415

Gln Pro Ser Gly Val Thr Ala Leu Thr Pro Arg Arg Asp Gly Leu Glu

420

425

430

15 Val Val Val Thr Pro Gly Ala Met Gly Ala Tyr Ala Cys Glu Cys Gln

435

440

445

Glu Gly Gly Ala Ala His Val Val Ala Ala Tyr Ser Leu Val Trp Gly

450

455

460

Ser Gln Arg Asp Ala Pro Ser Arg Ala His Thr Val Gly Ala Gly Leu

20 465

470

475

480

Ala Gly Phe Phe Leu Gly Ile Leu Ala Ala Ser Leu Thr Leu Ile Leu

485

490

495

Ile Gly Arg Arg Gln Gln Arg Arg Arg Gln Arg Glu Leu Leu Ala Arg

500

505

510

25 Asp Lys Val Gly Leu Asp Leu Gly Ala Pro Pro Ser Gly Thr Thr Ser

09284180-060999

515 520 525  
Tyr Ser Gln Asp Pro Pro Ser Pro Ser Pro Glu Asp Glu Arg Leu Pro  
530 535 540  
Leu Ala Leu Ala Lys Arg Gly Ser Gly Phe Gly Gly Phe Ser Pro Pro  
5 545 550 555 560  
Phe Leu Leu Asp Pro Cys Pro Ser Pro Ala His Ile Arg Leu Thr Gly  
565 570 575  
Ala Pro Leu Ala Thr Cys Asp Glu Thr Ser Ile  
580 585

10

SEQ ID NO: 7

SEQUENCE LENGTH: 196 base pairs

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

15

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL: No

ANTI-SENSE: No

ORIGINAL SOURCE:

20

ORGANISM: human (Homo sapiens)

TISSUE TYPE: brain

FEATURE:

FEATURE KEY: CDS

LOCATION: 1..196

25

IDENTIFICATION METHOD: E

09284180.060999

SEQUENCE DESCRIPTION:

AAATTGTACC ACAGCTGGCT CCTGGTTGGC TCCCGTACTG AGGTGACACA AGTGAATACA 60  
ACCAACTGTG GCCGTCTCCA GAGCTGCTCA GAGTGCATCC TGGCCCAGGA CCCAGTCTGT 120  
GCCTGGAGCT TCCGGCTGGA TGAGTGTGTG GCCCATGCCG GGGAGCACCG AGGGTTGGTC 180  
5 CAAGACATAG AGTCAG 196

SEQ ID NO: 8

SEQUENCE LENGTH: 30 base pairs

SEQUENCE TYPE: nucleic acid

10 STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Other nucleic acid, synthetic DNA

HYPOTHETICAL: No

ANTI-SENSE: No

15 SEQUENCE DESCRIPTION:

GATAAGGATC CGGGTCGCCG TCAGCAGCGT 30

SEQ ID NO: 9

SEQUENCE LENGTH: 27 base pairs

20 SEQUENCE TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Other nucleic acid, synthetic DNA

HYPOTHETICAL: No

25 ANTI-SENSE: Yes

09284180-050999

Q

SEQUENCE DESCRIPTION:

GGCTGGAATT CATTTCCCC GGCTTTA

27

SEQ ID NO: 10

5 SEQUENCE LENGTH: 333 base pairs

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

10 HYPOTHETICAL: No

ANTI-SENSE: No

ORIGINAL SOURCE:

ORGANISM: human (Homo sapiens)

TISSUE TYPE: brain

15 FEATURE:

FEATURE KEY: CDS

LOCATION: 1..333

IDENTIFICATION METHOD: E

SEQUENCE DESCRIPTION:

20 CCCCCGCCGG GTCCCGGGCA GCCTACAGCC TCGCCCTTCC CGCTACTGCT GCTGGCGGTG 60  
CTGAGCGGCC CGGTATCCGG CCGCGTCCCC CGCTCGGTGC CCAGAACCTC GCTTCCAATC 120  
TCTGAGGCTG ACTTCTGTCT CACCGGTTC GCAGTCCCTC ACACATACAA TTA CTCTGTT 180  
CTCCTTGTGG ATCCTGCCTC CCACACACTT TATGTTGGCG CCCGGGACAC CATCTTCGCT 240  
TTATCCCTGC CTTTCTCAGG GGAGAGACCC CGCAGGATTG ACTGGATGGT TCCTGAGGCT 300  
25 CACAGACAGA ACTGTAGGAA GAAAGGCAAG AAA 333

a

09284130 060999

SEQ ID NO: 11

SEQUENCE LENGTH: 111 amino acids

SEQUENCE TYPE: amino acid

5 TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: human (Homo sapiens)

TISSUE TYPE: brain

10 FEATURE:

FEATURE KEY: peptide

LOCATION: 1..111

IDENTIFICATION METHOD: F

SEQUENCE DESCRIPTION:

15 Pro Arg Pro Gly Pro Gly Gln Pro Thr Ala Ser Pro Phe Pro Leu Leu

5 10 15

Leu Leu Ala Val Leu Ser Gly Pro Val Ser Gly Arg Val Pro Arg Ser

20 25 30

Val Pro Arg Thr Ser Leu Pro Ile Ser Glu Ala Asp Phe Cys Leu Thr

20 35 40 45

Arg Phe Ala Val Pro His Thr Tyr Asn Tyr Ser Val Leu Leu Val Asp

50 55 60

Pro Ala Ser His Thr Leu Tyr Val Gly Ala Arg Asp Thr Ile Phe Ala

65 70 75 80

25 Leu Ser Leu Pro Phe Ser Gly Glu Arg Pro Arg Arg Ile Asp Trp Met

09084180 060999

Q

85

90

95

Val Pro Glu Ala His Arg Gln Asn Cys Arg Lys Lys Gly Lys Lys

100

105

110

a

00284180-060999

a

# SEQUENCE LISTING

<110> KIMURA, Toru  
KIKUCHI, Kaoru

<120> NOVEL SEMAPHORIN GENE: SEMAPHORIN W

<130> 20-4546P

<140>

<141>

<160> 14

<170> PatentIn Ver. 2.0

<210> 1

<211> 4008

<212> DNA

<213> Rattus norvegicus

<220>

<221> 5'UTR

<222> (1)..(75)

<220>

<223> Coding region from residue 76 to 2406

<220>

<221> 3'UTR

<222> ()..(3977)

<220>

<221> polyA\_signal

<222> ()..(4008)

<400> 1

```
gccgaggccc gcgcagtagc ggtactaagt agaggctgct ggacgcgccc caccggcac 60
caggcggagc cagagatgct tgccagggcc gagcggcccc gcccgggccc ccggccgcct 120
ccggtctttc ccttcccgcc gccgtgctcg ctgctgctgc tgctggcgat actaagcgcc 180
ccgggtgtgcg gccgcgtccc ccgctcagtg cccagaacct cgctgcccac ctccgaggct 240
gactcctatc tcacccgggt tgcagcgtct catacgtaca attactctgc tctccttgctg 300
gatcctgcct cccacacact ttacgtcggg gcacgggata gcatcttcgc tttaaccctc 360
cccttctctg gggaaagacc ccgaaggatc gactggatgg tacctgagac tcacagacag 420
aactgcagga agaaaggcaa gaaagaggac gaatgtcaca attttatcca gattctcgcc 480
attgtcaatg cctctcacct cctcacgtgc ggcaccttcg cttttgatcc gaagtgcggg 540
gttattgatg tgtccagttt ccagcaggtt gaaagacttg agagcggccg ggggaaatgt 600
ccttttgagc cagctcaacg gtcagcagct gtaatggctg ggggcgtcct ctacaccgcc 660
```



actgtgaaga	acttcctggg	gactgagccc	atcatctccc	gagctgtggg	tcgagctgag	720
gactggattc	gaacagagac	cttgtcatcc	tggcttaatg	ctccagcctt	tgtcgcagct	780
atggctcctga	gcccagctga	gtggggggat	gaagatggag	acgatgaaat	cttttttttc	840
ttcacggaga	cctcccaggt	gttggaactcc	tatgagcgca	tcaaggtccc	aagagtggcc	900
cgagtgtgtg	cgggggacct	tgggggcagg	aagacccttc	agcagagatg	gacgacgttt	960
ctgaaggctg	acctgctgtg	cccagggccc	gagcatggcc	gggcctccgg	ggttctgcag	1020
gctatggcag	agcttcggcc	tcagcctgga	gcgggaaccc	ccatctttta	tgggatcttt	1080
tcttcccagt	gggaaggagc	tgccatctct	gctgtgtgtg	ccttccgacc	ccaagacatc	1140
cgggcagtg	tgaatgggtcc	ctttagagag	ctaaaacatg	actgcaacag	gggactgcct	1200
gtcatggaca	acgaggtgcc	ccagcccaga	cctggagagt	gcacgcgcaa	caacatgaag	1260
ctccagcagt	ttggatcctc	actctccctg	ccagaccgcg	tgctcacctt	tatcagagac	1320
cacctctca	tggacaggcc	cgtgttcccc	gctgacggcc	gccccctgct	ggtcactaca	1380
gatacagcct	atctcagagt	cgtggcccac	agggtgacca	gcctctcagg	gaaagaatat	1440
gacgtgctct	acctggggac	agaggatgga	cacctccacc	gggctgtgcg	cattggagct	1500
cagctcagtg	tcttggagga	tctggccttg	ttcccagaac	cacagccggt	tgagagcatg	1560
aaattgtacc	acgattggct	cctgggtggc	tcccatactg	aggtgacaca	agtgaacacc	1620
agcaactgtg	gccgtctcca	gagctgctcg	gagtgtatcc	tggcccagga	ccccgtgtgc	1680
gcctggagct	tccggcttga	tgcttgtgtg	gcccacgccg	gcgagcaccg	cgggatgggt	1740
caagatatag	agtcagcgga	tgtctcttct	ttgtgtccaa	aagaacctgg	agaacatccc	1800
gtagtgtttg	aagttccggt	ggctactgtg	ggccacgtgg	tcctgccatg	ttcccccagt	1860
tctgcctggg	catcctgtgt	gtggcaccag	cccagtggag	tgactgcgct	cactccccgg	1920
agggatggac	tagagggtgt	ggtgacccca	ggggccatgg	gggcttatgc	ttgcgagtgt	1980
caggaggggtg	gagccgcccc	cgtggtggct	gcttatagct	tgggtgtggg	cagccagcgg	2040
ggaccctcaa	accgggcccc	caccgttgtg	ggggctggat	tggttggctt	tctcctgggt	2100
gttcttgcag	catccctcac	tctcctcctg	attggtcgcc	gtcagcagcg	tcggcgacag	2160
agggagcttc	tagctagaga	caagggtggc	ttagatctgg	gggctccacc	ttctgggacc	2220
acaagctata	gtcaggaccc	tccctctcct	tcgcctgaag	atgaacggct	gcccctggcc	2280
ctgggtaagc	ggggcagtg	ttttgggtgg	ttccctccac	ccttctgct	ggattcttgc	2340
ccaagcccag	cccacatccg	gctcactggg	gcgcctctag	ccacgtgtga	tgagacctcc	2400
atctaaagcc	ggggaaaatg	actgccagcc	atgagcagtc	tctggaacta	gtggctacca	2460
agaccatgat	catggctgct	cctttctctt	ggagtctgtg	tgttcacaca	ttagtgtctg	2520
tctctggac	ctggacctgg	cctttgcccc	gattcctgat	tctcatgaga	gatcaaccct	2580
gtaaccttct	gcgatggcct	cttgtcttgg	gcccatacag	ttgtggggtg	gagtaaggac	2640
ataggccccg	gaaagggaat	cagtgtggag	gtagtgtggg	cgtgtgtgcc	ctgcgtcctt	2700
gtggtggctg	tatgatattc	cagtctgctg	actctgggga	gcgcatgac	ccctgactgc	2760
cttgagatct	ctcccaactc	agtttccccct	tgctctggaa	gagtgtgtgt	ctataactg	2820
gtgtgcctag	aaggcctgtc	catgtgtgca	tggacgacag	ggccggtgcc	tcgggtgcttt	2880
tggggagtcg	gagagaaaag	ttggaatggg	ggacaactta	accctcggtg	gccagtgagg	2940
gaaaccacat	gcccgtcccc	atcacccccac	agcgcttctt	taactttgag	caaagttccc	3000
aaagtgacct	tctgggtggg	aagggcagca	ggacatgtgg	cccccgctct	tctccttgct	3060
tttcccttct	ggctgccaac	cactgccgtg	ccaccgctgc	gctttccctg	gctggagtgg	3120
aggctgagtc	ctctgtcctt	ggtttccatt	taaaatgaac	ttcacaacat	tctaaatatt	3180
gggggatgac	aaatgacttt	tttccccaga	aaagtgtgta	ggaaatacaa	gcagggttaa	3240
gaagatttgc	ctcagtgaact	ttcacccttg	ccctaaagca	ggagtccctc	agctagcgtc	3300
tgtggactcc	ctgaaattgt	atgcgtctgt	ggactccctg	aaattgtatg	caaagtgtct	3360
gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtttgcgtgc	atgtgtgcat	3420
gtgtgtttga	tggctttcat	cagattctca	aggccttaat	gaggttaaag	gaccacggcc	3480
tatagtcacc	acacttgggc	cacatggagg	aggtgttgct	ctctgaggca	gttccctccct	3540

```

ggcctgacctg aggccagccc ctggacacat tgctgctgga gacccacat ctctccagaa 3600
cttgaagct aggcctctgcg cgtgcttgaa ggcaccacca tctcccttct tgcttcattc 3660
tcctgtgtgc tctgcctctg ctgagtcctg ctcttggcct gtgaatgtgc ctgcccgtc 3720
cctggtgggg gacctcaaac cccagtgtg atgctaccct ttccagtggg agtttctgtt 3780
ctgctttcct tgacagcagc ctgtgaacta ctcacgagtc cccttggttt ggagttcccg 3840
gtggctttga gtaggatctt tggcgtggca tctaacctag cagcattgat cgttcattgt 3900
aaagtgggga tatacctacc tcagggttgc tgcaaggatc aaacgaggaa acgtataaat 3960
aaagcattac ccacagcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4008

```

<210> 2

<211> 2331

<212> DNA

<213> *Rattus norvegicus*

<220>

<223> Coding region from residue 1 to 2331

<400> 2

```

atgcttgcca gggccgagcg gccccgcccg ggcccccggc cgcctccggt ctttcccttc 60
ccgccgcgcg tgtcgtgct gctgctgctg gcgatactaa gcgccccggt gtgcggccgc 120
gtcccccgct cagtgcccg aacctcgctg cccatctccg aggcctgactc ctatctcacc 180
cggtttgca cgtctcatal gtacaattac tctgctctcc ttgtggatcc tgcctccac 240
acactttacg tcggtgcacg ggatagcatc ttgcctttaa ccctcccctt ctctggggaa 300
agaccccgaa ggatcgactg gatggtacct gagactcaca gacagaactg caggaagaaa 360
ggcaagaaag aggcagaaatg tcacaatttt atccagattc tcgccattgt caatgcctct 420
cacctcctca cgtgcggcac cttcgctttt gatccgaagt gcggggttat tgatgtgtcc 480
agtttccagc aggttgaaag acttgagagc ggccggggga aatgtccttt tgagccagct 540
caacggtcag cagctgtaat ggctgggggc gtcctctaca ccgccactgt gaagaacttc 600
ctggggactg agcccatcat ctcccgagct gtgggtcgag ctgaggactg gattcgaaca 660
gagaccttgt catcctggct taatgctcca gcctttgtcg cagctatggt cctgagccca 720
gctgagtggg gggatgaaga tggagacgat gaaatctttt ttttcttcac ggagacctcc 780
cgagtgttg actcctatga gcgcataag gtcccaagag tggcccgagt gtgtgcgggg 840
gaccttgggg gcaggaagac ccttcagcag agatggacga cgtttctgaa ggctgacctg 900
ctgtgcccag ggcccagca tggccgggccc tccggggttc tgcaggctat ggcagagctt 960
cggcctcagc ctggagcggg aacccccatc ttttatggga tcttttctc ccagtgggaa 1020
ggagctgcca tctctgctgt gtgtgccttc cgacccaag acatccgggc agtgctgaat 1080
ggtcccttta gagagctaaa acatgactgc aacaggggac tgcctgtcat ggacaacgag 1140
gtgccccagc ccagacctgg agagtgcac gccaacaaca tgaagctcca gcagtttgga 1200
tcctcactct cctgcccaga ccgcgtgctc acctttatca gagaccacc tctcatggac 1260
aggcccggtg tcccggctga cggccgcccc ctgctggtca ctacagatac agcctatctc 1320
agagtcgtgg cccacagggt gaccagcctc tcagggaaag aatatgacgt gctctacctg 1380
gggacagagg atggacacct ccaccgggct gtgcgcattg gagctcagct cagtgtcttg 1440
gaggatctgg ccttgttccc agaaccacag ccggttgaga gcatgaaatt gtaccacgat 1500
tggctcctgg tgggtccca tactgaggtg acacaagtga acaccagcaa ctgtggccgt 1560
ctccagagct gctcggagtg tatcctggcc caggacccc tgtgcgcctg gagcttccgg 1620
cttgatgctt gtgtggccca cggcggcgag caccgcggga tggttcaaga tatagagtca 1680
gcggatgtct cttctttgtg tccaaaagaa cctggagaac atcccgtagt gtttgaagtt 1740
ccggtggcta ctgtgggcca cgtggtcctg ccatgttccc ccagttctgc ctgggcatcc 1800

```

tgtgtgtggc accagcccag tggagtgact gcgctcactc cccggaggga tggactagag 1860  
 gtggtggtga cccaggggc catgggggt tatgcttgcg agtgtcagga ggggtggagcc 1920  
 gcccggtgg tggctgctta tagcttggtg tggggcagcc agcggggacc ctcaaaccgg 1980  
 gccacaccg ttgtgggggc tggattggtt ggctttctcc tgggtgttct tgcagcatcc 2040  
 ctactctcc tctgattgg tcgccgtcag cagcgtcggc gacagaggga gcttctagct 2100  
 agagacaagg tgggcttaga tctgggggct ccaccttctg ggaccacaag ctatagtcag 2160  
 gacctccct ctcttcgcc tgaagatgaa cggctgcccc tggccctggg taagcggggc 2220  
 agtggttttg gtggcttccc tccaccttc ctgctggatt cttgcccaag cccagcccac 2280  
 atccggtca ctggggcgcc tctagccacg tgtgatgaga cctccatcta a 2331

<210> 3  
 <211> 776  
 <212> PRT  
 <213> Rattus norvegicus

<400> 3  
 Met Leu Ala Arg Ala Glu Arg Pro Arg Pro Gly Pro Arg Pro Pro Pro  
 1 5 10 15  
 Val Phe Pro Phe Pro Pro Pro Leu Ser Leu Leu Leu Leu Leu Ala Ile  
 20 25 30  
 Leu Ser Ala Pro Val Cys Gly Arg Val Pro Arg Ser Val Pro Arg Thr  
 35 40 45  
 Ser Leu Pro Ile Ser Glu Ala Asp Ser Tyr Leu Thr Arg Phe Ala Ala  
 50 55 60  
 Ser His Thr Tyr Asn Tyr Ser Ala Leu Leu Val Asp Pro Ala Ser His  
 65 70 75 80  
 Thr Leu Tyr Val Gly Ala Arg Asp Ser Ile Phe Ala Leu Thr Leu Pro  
 85 90 95  
 Phe Ser Gly Glu Arg Pro Arg Arg Ile Asp Trp Met Val Pro Glu Thr  
 100 105 110  
 His Arg Gln Asn Cys Arg Lys Lys Gly Lys Lys Glu Asp Glu Cys His  
 115 120 125  
 Asn Phe Ile Gln Ile Leu Ala Ile Val Asn Ala Ser His Leu Leu Thr  
 130 135 140  
 Cys Gly Thr Phe Ala Phe Asp Pro Lys Cys Gly Val Ile Asp Val Ser  
 145 150 155 160  
 Ser Phe Gln Gln Val Glu Arg Leu Glu Ser Gly Arg Gly Lys Cys Pro  
 165 170 175

Phe Glu Pro Ala Gln Arg Ser Ala Ala Val Met Ala Gly Gly Val Leu  
180 185 190  
Tyr Thr Ala Thr Val Lys Asn Phe Leu Gly Thr Glu Pro Ile Ile Ser  
195 200 205  
Arg Ala Val Gly Arg Ala Glu Asp Trp Ile Arg Thr Glu Thr Leu Ser  
210 215 220  
Ser Trp Leu Asn Ala Pro Ala Phe Val Ala Ala Met Val Leu Ser Pro  
225 230 235 240  
Ala Glu Trp Gly Asp Glu Asp Gly Asp Asp Glu Ile Phe Phe Phe Phe  
245 250 255  
Thr Glu Thr Ser Arg Val Leu Asp Ser Tyr Glu Arg Ile Lys Val Pro  
260 265 270  
Arg Val Ala Arg Val Cys Ala Gly Asp Leu Gly Gly Arg Lys Thr Leu  
275 280 285  
Gln Gln Arg Trp Thr Thr Phe Leu Lys Ala Asp Leu Leu Cys Pro Gly  
290 295 300  
Pro Glu His Gly Arg Ala Ser Gly Val Leu Gln Ala Met Ala Glu Leu  
305 310 315 320  
Arg Pro Gln Pro Gly Ala Gly Thr Pro Ile Phe Tyr Gly Ile Phe Ser  
325 330 335  
Ser Gln Trp Glu Gly Ala Ala Ile Ser Ala Val Cys Ala Phe Arg Pro  
340 345 350  
Gln Asp Ile Arg Ala Val Leu Asn Gly Pro Phe Arg Glu Leu Lys His  
355 360 365  
Asp Cys Asn Arg Gly Leu Pro Val Met Asp Asn Glu Val Pro Gln Pro  
370 375 380  
Arg Pro Gly Glu Cys Ile Ala Asn Asn Met Lys Leu Gln Gln Phe Gly  
385 390 395 400  
Ser Ser Leu Ser Leu Pro Asp Arg Val Leu Thr Phe Ile Arg Asp His  
405 410 415  
Pro Leu Met Asp Arg Pro Val Phe Pro Ala Asp Gly Arg Pro Leu Leu  
420 425 430

Val Thr Thr Asp Thr Ala Tyr Leu Arg Val Val Ala His Arg Val Thr  
 435 440 445

Ser Leu Ser Gly Lys Glu Tyr Asp Val Leu Tyr Leu Gly Thr Glu Asp  
 450 455 460

Gly His Leu His Arg Ala Val Arg Ile Gly Ala Gln Leu Ser Val Leu  
 465 470 475 480

Glu Asp Leu Ala Leu Phe Pro Glu Pro Gln Pro Val Glu Ser Met Lys  
 485 490 495

Leu Tyr His Asp Trp Leu Leu Val Gly Ser His Thr Glu Val Thr Gln  
 500 505 510

Val Asn Thr Ser Asn Cys Gly Arg Leu Gln Ser Cys Ser Glu Cys Ile  
 515 520 525

Leu Ala Gln Asp Pro Val Cys Ala Trp Ser Phe Arg Leu Asp Ala Cys  
 530 535 540

Val Ala His Ala Gly Glu His Arg Gly Met Val Gln Asp Ile Glu Ser  
 545 550 555 560

Ala Asp Val Ser Ser Leu Cys Pro Lys Glu Pro Gly Glu His Pro Val  
 565 570 575

Val Phe Glu Val Pro Val Ala Thr Val Gly His Val Val Leu Pro Cys  
 580 585 590

Ser Pro Ser Ser Ala Trp Ala Ser Cys Val Trp His Gln Pro Ser Gly  
 595 600 605

Val Thr Ala Leu Thr Pro Arg Arg Asp Gly Leu Glu Val Val Val Thr  
 610 615 620

Pro Gly Ala Met Gly Ala Tyr Ala Cys Glu Cys Gln Glu Gly Gly Ala  
 625 630 635 640

Ala Arg Val Val Ala Ala Tyr Ser Leu Val Trp Gly Ser Gln Arg Gly  
 645 650 655

Pro Ser Asn Arg Ala His Thr Val Val Gly Ala Gly Leu Val Gly Phe  
 660 665 670

Leu Leu Gly Val Leu Ala Ala Ser Leu Thr Leu Leu Leu Ile Gly Arg  
 675 680 685

Arg Gln Gln Arg Arg Arg Gln Arg Glu Leu Leu Ala Arg Asp Lys Val  
690 695 700

Gly Leu Asp Leu Gly Ala Pro Pro Ser Gly Thr Thr Ser Tyr Ser Gln  
705 710 715 720

Asp Pro Pro Ser Pro Ser Pro Glu Asp Glu Arg Leu Pro Leu Ala Leu  
725 730 735

Gly Lys Arg Gly Ser Gly Phe Gly Gly Phe Pro Pro Pro Phe Leu Leu  
740 745 750

Asp Ser Cys Pro Ser Pro Ala His Ile Arg Leu Thr Gly Ala Pro Leu  
755 760 765

Ala Thr Cys Asp Glu Thr Ser Ile  
770 775

<210> 4

<211> 2315

<212> DNA

<213> Homo sapiens

<220>

<223> Coding region from residue 1 to 1764

<220>

<221> 3'UTR

<222> ()..(2315)

<400> 4

gggggtgtcc tctatgctgc cactgtgaaa aactacctgg ggacggagcc aattatcacc 60  
agagcagtgg gtcgtgccga ggactggatt cggacagata ccttgccttc ctggctgaac 120  
gccccagcct ttgtcgcagc cgtggccttg agcccagccg aatgggggga tgaagatgga 180  
gacgacgaaa tctacttctt ctttacggag acttcccag catttgactc atacgagcgc 240  
attaaagtcc cacgggtggc ccgtgtgtgt gcgggggacc tcggggggcc gaagaccctc 300  
cagcagagat ggacgacgtt tttgaaagct gacctgctct gtccagggcc tgagcatggc 360  
cgggcctcca gtgtcctgca ggatgttgct gtgcttcgac ctgagcttgg ggcagggact 420  
cccatctttt atggcatctt ttcttcccag tgggaggggg ctactatctc tgctgtctgt 480  
gccttccgac cacaagacat tcggacagtg ctgaatggtc ccttcagaga actaaaacat 540  
gactgcaaca gaggactgcc tgtcgtggac aatgatgtgc cccagcccag acctggagag 600  
tgcattacca acaacatgaa gctccggcac tttggctcat ctctctccct gcctgaccgc 660  
gtactcacct tcatccggga ccaccactc atggacaggc cagtgtttcc agctgatggc 720  
caccctctgc tggctactac agatacagcc tatctcagag tcgtggccca cagggtgacc 780  
agcctctcag ggaaagagta tgatgtgctc tacctgggga cagaggatgg acacctccac 840  
cgagcagtgc ggatcggagc tcagctcagc gttcttgaag atctggcctt attcccagag 900

```

ccacagccag ttgagaacat gaaattgtac cacagctggc tcctggttgg ctcccgtact 960
gaggtgacac aagtgaatac aaccaactgt ggccgtctcc agagctgctc agagtgcac 1020
ctggcccagg acccagtctg tgccctggagc ttccggctgg atgagtgtgt ggcccatgcc 1080
ggggagcacc gagggtttgt ccaagacata gagtgcagc atgtctcctc tttgtgtcct 1140
aaagagcctg gagaacgtcc agtagtgttt gaagttcccg tggctacagc tgcgcatgtg 1200
gtcttgccat gttctccaag ctacagcatgg gcatcctgtg tgtggcacca gccagtgga 1260
gtgactgcac tcaccccccg gcgggatgga ctggagggtg tggtgacccc aggggccaatg 1320
ggcgcttatg cctgtgaatg tcaggagggt ggggcagccc atgtggtagc agcttacagc 1380
ttggtatggg gcagccagcg agatgctccg agccgggccc acacagtggg ggcgggactg 1440
gctggcttct tcttggggat tctgcagca tccctgactc tcattctgat tggtcggcgt 1500
cagcagcgac ggcgacagag ggaacttctg gctagagaca aggtgggcct ggacctggg 1560
gctccacctt ctgggaccac aagctacagc caagaccctc cctccccctc tcctgaagat 1620
gagcggttgc cgctggccct ggccaagagg ggcagtggct ttggtggatt ctaccacccc 1680
ttcctgcttg atccttgccc aagcccagcc cacattcggc taactggggc tcctctagcc 1740
acatgtgatg aaacatccat ctagagctgg gcaaagacc actagtgtat aagtgtcac 1800
tggaacggag tgaccactga gatgctgggg gtcactgggc ctggaagacc atcccagcct 1860
ctgagttctc tttgagtatg agtgattact tggatttttag tatctgttct ctctgagcct 1920
ggatgggctt ggggcccagac ctttgctga ttcctgattc ccatgagaaa tcagaactgc 1980
tttctgcagc aaatcagggc ttccccctaa catctgaact cctgtaaacc ttcattccctg 2040
gccccctatc ttgggcccct tagttttggg gatggggcac agggcatagc tatgactttg 2100
ctttctgggt ggagcctggc cggaaggaag agccctggag gtggttgggg gcaaagtgtc 2160
cctgagtcct tgggggtggt ctgcttattc ttcaagttta tctgaatctg tggggagtgc 2220
atgatcccca tgttgcaata tggagtctct gccctgagat cttccccatc tcagttttcc 2280
ttccatgaaa gagtacgtgt aaatacatag tgttc 2315

```

<210> 5

<211> 1761

<212> DNA

<213> Homo sapiens

<220>

<223> Coding region from residue 1 to 1761

<400> 5

```

gggggtgtcc tctatgctgc cactgtgaaa aactacctgg ggacggagcc aattatcacc 60
agagcagtgg gtcgtgccga ggactggatt cggacagata ccttgccctc ctggctgaac 120
gccccagcct ttgtgcgagc cgtggccttg agcccagccg aatgggggga tgaagatgga 180
gacgacgaaa tctacttctt ctttacggag acttcccag cathttgactc atacgagcgc 240
attaaagtcc cacgggtggc ccgtgtgtgt gcgggggacc tcggggggccg gaagaccctc 300
cagcagagat ggacgacgtt tttgaaaagt gacctgctct gtccagggcc tgagcatggc 360
cgggcctcca gtgtcctgca ggatgttgct gtgcttcgac ctgagcttgg ggcagggact 420
cccatctttt atggcatctt ttcttcccag tgggaggggg ctactatctc tgctgtctgt 480
gccttccgac cacaagacat tcggacagtg ctgaatggtc ccttcagaga actaaaacat 540
gactgcaaca gaggactgcc tgtcgtggac aatgatgtgc ccagcccag acctggagag 600
tgcataacca acaacatgaa gctccggcac tttggctcat ctctctccct gcctgaccgc 660
gtactcacct tcatccggga ccaccactc atggacaggc cagtgtttcc agctgatggc 720
caccacctgc tggctactac agatacagcc tatctcagag tcgtggccca cagggtgacc 780
agcctctcag ggaaagagta tgatgtgctc tacctgggga cagaggatgg acacctccac 840

```

cgagcagtgc ggatcggagc tcagctcagc gttcttgaag atctggcctt attcccagag 900  
 ccacagccag ttgagaacat gaaattgtac cacagctggc tcctgggttg cccccgtact 960  
 gaggtgacac aagtgaatac aaccaactgt ggccgtctcc agagctgctc agagtgcac 1020  
 ctggcccagg acccagtctg tgccctggagc ttccggctgg atgagtgtgt ggcccatgcc 1080  
 ggggagcacc gagggtttgt ccaagacata gagtcagcag atgtctcctc tttgtgtcct 1140  
 aaagagcctg gagaacgtcc agtagtgttt gaagttcccg tggctacagc tgcgcatgtg 1200  
 gtcttgccat gttctccaag ctacagcatgg gcatcctgtg tgtggcacca gcccagtggg 1260  
 gtgactgcac tcaccccccg gcgggatgga ctggaggtgg tggtgacccc agggggccatg 1320  
 ggcgcttatg cctgtgaatg tcaggagggt ggggcagccc atgtggtagc agcttacagc 1380  
 ttggtatggg gcagccagcg agatgctccg agccgggccc acacagtggg ggcgggactg 1440  
 gctggcttct tcttggggat tctgcagca tccctgactc tcattctgat tggtcggcgt 1500  
 cagcagcgac ggcgacagag ggaacttctg gctagagaca aggtgggcct ggacctgggg 1560  
 gctccacctt ctgggaccac aagctacagc caagaccctc cctccccctc tcctgaagat 1620  
 gagcggttgc cgctggccct ggccaagagg ggcagtggct ttggtggatt ctaccacccc 1680  
 ttctgcttg atccttgccc aagcccagcc cacattcggc taactggggc tcctctagcc 1740  
 acatgtgatg aaacatccat c 1761

<210> 6  
 <211> 587  
 <212> PRT  
 <213> Homo sapiens

<400> 6  
 Gly Gly Val Leu Tyr Ala Ala Thr Val Lys Asn Tyr Leu Gly Thr Glu  
 1 5 10 15  
 Pro Ile Ile Thr Arg Ala Val Gly Arg Ala Glu Asp Trp Ile Arg Thr  
 20 25 30  
 Asp Thr Leu Pro Ser Trp Leu Asn Ala Pro Ala Phe Val Ala Ala Val  
 35 40 45  
 Ala Leu Ser Pro Ala Glu Trp Gly Asp Glu Asp Gly Asp Asp Glu Ile  
 50 55 60  
 Tyr Phe Phe Phe Thr Glu Thr Ser Arg Ala Phe Asp Ser Tyr Glu Arg  
 65 70 75 80  
 Ile Lys Val Pro Arg Val Ala Arg Val Cys Ala Gly Asp Leu Gly Gly  
 85 90 95  
 Arg Lys Thr Leu Gln Gln Arg Trp Thr Thr Phe Leu Lys Ala Asp Leu  
 100 105 110  
 Leu Cys Pro Gly Pro Glu His Gly Arg Ala Ser Ser Val Leu Gln Asp  
 115 120 125  
 Val Ala Val Leu Arg Pro Glu Leu Gly Ala Gly Thr Pro Ile Phe Tyr



130	135	140
Gly Ile Phe Ser Ser Gln Trp Glu Gly Ala Thr Ile Ser Ala Val Cys		
145	150	155 160
Ala Phe Arg Pro Gln Asp Ile Arg Thr Val Leu Asn Gly Pro Phe Arg		
	165	170 175
Glu Leu Lys His Asp Cys Asn Arg Gly Leu Pro Val Val Asp Asn Asp		
	180	185 190
Val Pro Gln Pro Arg Pro Gly Glu Cys Ile Thr Asn Asn Met Lys Leu		
	195	200 205
Arg His Phe Gly Ser Ser Leu Ser Leu Pro Asp Arg Val Leu Thr Phe		
	210	215 220
Ile Arg Asp His Pro Leu Met Asp Arg Pro Val Phe Pro Ala Asp Gly		
	225	230 235 240
His Pro Leu Leu Val Thr Thr Asp Thr Ala Tyr Leu Arg Val Val Ala		
	245	250 255
His Arg Val Thr Ser Leu Ser Gly Lys Glu Tyr Asp Val Leu Tyr Leu		
	260	265 270
Gly Thr Glu Asp Gly His Leu His Arg Ala Val Arg Ile Gly Ala Gln		
	275	280 285
Leu Ser Val Leu Glu Asp Leu Ala Leu Phe Pro Glu Pro Gln Pro Val		
	290	295 300
Glu Asn Met Lys Leu Tyr His Ser Trp Leu Leu Val Gly Ser Arg Thr		
	305	310 315 320
Glu Val Thr Gln Val Asn Thr Thr Asn Cys Gly Arg Leu Gln Ser Cys		
	325	330 335
Ser Glu Cys Ile Leu Ala Gln Asp Pro Val Cys Ala Trp Ser Phe Arg		
	340	345 350
Leu Asp Glu Cys Val Ala His Ala Gly Glu His Arg Gly Leu Val Gln		
	355	360 365
Asp Ile Glu Ser Ala Asp Val Ser Ser Leu Cys Pro Lys Glu Pro Gly		
	370	375 380
Glu Arg Pro Val Val Phe Glu Val Pro Val Ala Thr Ala Ala His Val		

385		390		395		400									
Val	Leu	Pro	Cys	Ser	Pro	Ser	Ser	Ala	Trp	Ala	Ser	Cys	Val	Trp	His
				405					410					415	
Gln	Pro	Ser	Gly	Val	Thr	Ala	Leu	Thr	Pro	Arg	Arg	Asp	Gly	Leu	Glu
			420					425					430		
Val	Val	Val	Thr	Pro	Gly	Ala	Met	Gly	Ala	Tyr	Ala	Cys	Glu	Cys	Gln
			435				440					445			
Glu	Gly	Gly	Ala	Ala	His	Val	Val	Ala	Ala	Tyr	Ser	Leu	Val	Trp	Gly
	450					455					460				
Ser	Gln	Arg	Asp	Ala	Pro	Ser	Arg	Ala	His	Thr	Val	Gly	Ala	Gly	Leu
465					470					475					480
Ala	Gly	Phe	Phe	Leu	Gly	Ile	Leu	Ala	Ala	Ser	Leu	Thr	Leu	Ile	Leu
				485				490						495	
Ile	Gly	Arg	Arg	Gln	Gln	Arg	Arg	Arg	Gln	Arg	Glu	Leu	Leu	Ala	Arg
				500				505					510		
Asp	Lys	Val	Gly	Leu	Asp	Leu	Gly	Ala	Pro	Pro	Ser	Gly	Thr	Thr	Ser
			515				520					525			
Tyr	Ser	Gln	Asp	Pro	Pro	Ser	Pro	Ser	Pro	Glu	Asp	Glu	Arg	Leu	Pro
	530					535					540				
Leu	Ala	Leu	Ala	Lys	Arg	Gly	Ser	Gly	Phe	Gly	Gly	Phe	Ser	Pro	Pro
545					550					555					560
Phe	Leu	Leu	Asp	Pro	Cys	Pro	Ser	Pro	Ala	His	Ile	Arg	Leu	Thr	Gly
				565					570					575	
Ala	Pro	Leu	Ala	Thr	Cys	Asp	Glu	Thr	Ser	Ile					
			580					585							

<210> 7

<211> 196

<212> DNA

<213> Homo sapiens

<220>

<223> Coding region from residue 1 to 196

<400> 7

```

aaattgtacc acagctggct cctggttggc tcccgtactg aggtgacaca agtgaataca 60
accaactgtg gccgtctcca gagctgctca gagtgcaccc tggcccagga cccagtctgt 120
gcctggagct tccggctgga tgagtgtgtg gcccatgccg gggagcaccg agggttggtc 180
caagacatag agtcag 196

```

```

<210> 8
<211> 30
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic DNA

```

```

<400> 8
gataaggatc cgggtcgccg tcagcagcgt 30

```

```

<210> 9
<211> 27
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence: Synthetic DNA

```

```

<400> 9
ggctggaatt cattttcccc ggcttta 27

```

```

<210> 10
<211> 333
<212> DNA
<213> Homo sapiens

```

```

<220>
<223> Coding region from residue 1 to 333

```

```

<400> 10
ccccggccgg gtcccgggca gcctacagcc tcgcccttcc cgctactgct gctggcggtg 60
ctgagcggcc cggatatccg ccgcgtcccc cgctcggtgc ccagaacctc gttccaatc 120
tctgaggctg acttctgtct caccgggttc gcagtcctc acacatacaa ttactctgtt 180
ctccttgtgg atcctgcctc ccacacactt tatgttggcg cccgggacac catcttcgct 240
ttatccctgc cttctcagg ggagagaccc cgcaggattg actggatggt tcctgaggct 300
cacagacaga actgtaggaa gaaaggcaag aaa 333

```

```

<210> 11
<211> 111
<212> PRT
<213> Homo sapiens

```

<400> 11

Pro Arg Pro Gly Pro Gly Gln Pro Thr Ala Ser Pro Phe Pro Leu Leu  
1 5 10 15

Leu Leu Ala Val Leu Ser Gly Pro Val Ser Gly Arg Val Pro Arg Ser  
20 25 30

Val Pro Arg Thr Ser Leu Pro Ile Ser Glu Ala Asp Phe Cys Leu Thr  
35 40 45

Arg Phe Ala Val Pro His Thr Tyr Asn Tyr Ser Val Leu Leu Val Asp  
50 55 60

Pro Ala Ser His Thr Leu Tyr Val Gly Ala Arg Asp Thr Ile Phe Ala  
65 70 75 80

Leu Ser Leu Pro Phe Ser Gly Glu Arg Pro Arg Arg Ile Asp Trp Met  
85 90 95

Val Pro Glu Ala His Arg Gln Asn Cys Arg Lys Lys Gly Lys Lys  
100 105 110

<210> 12

<211> 7

<212> PRT

<213> Homo sapiens

<400> 12

Gln Asp Pro Val Cys Ala Trp  
1 5

<210> 13

<211> 6

<212> PRT

<213> Homo sapiens

<220>

<223> First Xaa = Gln or Arg

<220>

<223> Second Xaa = Ala or Gly

<400> 13

Xaa Asp Tyr Cys Xaa Trp  
1 5

<210> 14

<211> 14

<212> PRT

<213> Unknown

<220>

<223> Description of Unknown Organism: Myc tag

<400> 14

Asp Ile Gly Gly Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

1

5

10